Request to Archive
With The National Centers for Environmental Information
For Global Land Surface Temperature Databank
Provided by NCDC/IAB

2012-11-16

This information will be used by NCEI to conduct an appraisal and make a decision on the request.

1. Who is the primary point of contact for this request?

Jay Lawrimore

NCDC

Ingest and Analysis Branch Chief

+1 828-271-4750

jay.lawrimore@noaa.gov

2. Name the organization or group responsible for creating the dataset.

DOC/NOAA/NESDIS/NCDC > National Climatic Data Center, NESDIS, NOAA, U.S. Department of Commerce

3. Provide an overview summarizing the scope of data you want to archive. Describe the outputs, data variables, including their measurement resolution and coverage.

The global land surface temperature databank contains monthly timescale mean, max, and min temperature for approximately 40,000 stations globally. It was developed as part of the International Surface Temperature Initiative and released in Beta form in early October 2012. Following 3 months of user review and comments the dataset will be released as Version 1 in early calendar year 2013. The archive of the databank will need to include Stage 1, Stage 2, and Stage 3 data.

4. What is the time period covered by the dataset? (YYYY-MM-DD, YYYY-MM or YYYY)

From 1800-01-01

Ongoing as continuous updates to the data record

5. Edition or version number(s) of the dataset:

version 1

6. Describe the level to which the data are processed. For example, are these unprocessed raw observations, derived parameters, quality controlled or inter-calibrated data, etc.?

These are ISTI Stage 3 data, developed from raw observations collected from 45 or more sources provided by others throughout the international community. The data are combined into a merged dataset using processes approved by the ISTI Databank working group.

7. Approximate date when the dataset was or will be released to the public:

2013-02-01

8. Who are the expected users of the archived data? How will the archived data be used?

NCDC and it's international partners, e.g., UK Met Office, Australia BOM, JMO.

9. Has the dataset undergone user evaluation and/or an independent review process? Did NCEI participate in design reviews?

The ISTI Databank Working Group have reviewed the data. They are out in beta release and we are receiving feedback from the community. Also a journal article will be submitted in November.

10. Describe the dataset's relationship to other archived datasets, such as earlier versions or related source data. If this is a new version, how does it improve upon the previous version(s)?

This dataset will form the foundation for GHCN-Monthly version 4.

11. List the input datasets and ancillary information used to produce the data.

Daily and monthly timescale source datasets are documented in the Stage 1 directories on the ISTI Databank ftp sites: ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank/daily/stage1/ and ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank/monthly/stage1/

12. List web pages and other links that provide information on the data.

The metadata includes basic information such as station name (when available), location, and elevation (when available). Metadata also include the data provenance tracking flags in Stage 2.

13. List the kinds of documents, metadata and code that are available for archiving. For example, data format specifications, user guides, algorithm documentation, metadata compliant with a standard such as ISO 19115, source code, platform/instrument metadata, data/process flow diagrams, etc.

1. Data readme files for each data source at
   ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank/README_DATABANK_MAIN.
2. Software used for creating the Stage 3 product at
   ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank/monthly/stage3/recommended/code/

14. Indicate the data file format(s).

1. ASCII

15. Are the data files compressed?

gzip

16. Provide details on how the files are named and how they are organized (e.g., file_name_pattern_YYYYMM.tar in monthly aggregations).

Stage 3: recommended-ghcn_format.monthly.stage3.v1.0.0-v1.yearmoda.tar.gz
Stage 2: source.monthly.stage2.1.0.yearmoda.tar.gz
Stage 1: source.monthly.stage1.1.0.yearmoda.tar.gz

17. Explain how to access sample data files and/or a file listing for previewing. If it is not available now, when will it be available?

Accessible via ftp: ftp://ftp.ncdc.noaa.gov/pub/data/globaldatabank/

18. What is the total data volume to be submitted?

Historic Data: all historic data or data submitted as a completed collection.

Total Data Volume:	25GB
Number of Data Files:	1100

Continuous Data: data volume rate for a continuous data production.

Total Data Volume Rate: 1GB per Month
Data File Frequency: 100 per Month
Data Production Start:

19. Are later updates, revisions or replacement files anticipated? If so, explain the conditions for submitting these additional data to the archive.
Subsequent versions will be developed in the future. Timeline TBD.

20. Describe the server that will connect to the ingest server at NCEI for submitting the data.
Physical Location: Asheville
System Name: Global Databank
System Owner: GCAD/IAB
Additional Information: Add comments as needed on applicable data types, etc.

21. What are the possible methods for submitting the data to NCEI? Select all that apply.
1. FTP PUSH

22. Identify how you would like NCEI to distribute the data. Web access support depends on the resources available for the dataset.
1. Direct download links

23. Will there be any distribution, usage, or other restrictions that apply to the data in the archive?
No known constraints apply to the data.

24. Discuss the rationale for archiving the dataset and the anticipated benefits. Mention any risks associated with not archiving the dataset at NCEI.
This is the global repository for all monthly timescale land surface observations from the 1800s to present. It brings together data from more than 45 sources into a single integrated dataset. From this dataset GHCN-Monthly version 4 will be developed and many other applications are expected in the international community.

25. Are the data archived at another facility or are there plans to do so? Please explain.
No

26. Is there an existing agreement or requirement driving this request to archive? Have you already contacted someone at NCEI?
NCDC's commitment to development of GHCN-M version 4 and the International Surface Temperature Initiative are the requirements driving this effort.

27. Do you have a data management plan for your data?
No

28. Have funds been allocated to archive the data at NCEI?
No

29. Identify the affiliated research project, its sponsor, and any project/grant ID as applicable.
International Surface Temperature Initiative

30. Is there a desired deadline for NCEI to archive and provide access to the data?

Archive by: 2013-02-28

Accessible by:

31. Add any other pertinent information for this request.

None